



Session 11: 6 Nov 2018
SIRE503: IMB



Genetic Risk Profile

Bhoom Suktitipat, MD, PhD

Department of Biochemistry, Faculty of Medicine Siriraj Hospital,
Integrative Computational Bioscience Center

Mahidol University

bhoom.suk@mahidol.edu



Learning Objectives



- Can explain
 - what genetic risk profile (GRP) is
 - how to construct GRP/calculate GRP
- Can give examples how GRP is used
- Can calculate simple genetic score using weighted sum method.



Outline

- Measuring Risks
- Some terminology
- Rationale for GRP
- Framework for creating GRP
- Obtaining the reference for genetic risk
- PRS for risk stratification
- PRS and behavior



Measuring Risk



odds | ädz |

pl. noun

the ratio between the amounts staked by the parties to a bet, based on the expected probability either way: *the bookies are offering odds of 8-1 | it is possible for the race to be won at very long odds.*

- (usually **the odds**) the chances or likelihood of something happening or being the case: *the odds are that he is no longer alive | the odds against this ever happening are high.*
 - (usually **the odds**) superiority in strength, power, or resources; advantage: *she clung to the lead against all the odds | the odds were overwhelmingly in favor of the banks rather than the customer.*
-
- Odds (not the same as “odd” – without “s”)
Odd: different from what is usual



Odds Ratio

- Probability & odds
 - Odds = $P(Y = 1) / P(Y = 0) = p/(1-p)$
 - $Y \in \{0,1\}$
- Odds Ratio
 - Ratio of two odds
 - Odds of disease among cases divided Odds of disease among control

Usually use in case-control studies of rare diseases/events



Relative Risk

- Risk ratio: ratio of incidence in exposed vs unexposed group.

		Group	
		Experimental (E)	Control (C)
Events (E)	EE	EC	
Non-events (N)	NE	NC	

$$RR = \frac{\frac{EE}{(EE+NE)}}{\frac{EC}{(EC+NC)}} \approx \frac{EE \times NC}{NE \times EC} = OR$$

$EE \ll NE, EE + NE \approx NE$
 $EC \ll NC; EC + NC \approx NC$

Usually use in case-control studies i.e. clinical trial when events (treatment response) is quite common)



Hazard Ratio



- Hazard rate: instantaneous risk at a point in time

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{observed events in interval } [t, t + \Delta t] / N(t)}{\Delta t}$$

- Hazard ratio: a ratio of two hazards i.e. patients vs healthy individuals
- Usually use in survival analysis (longitudinal cohort)



Other Type of Risks

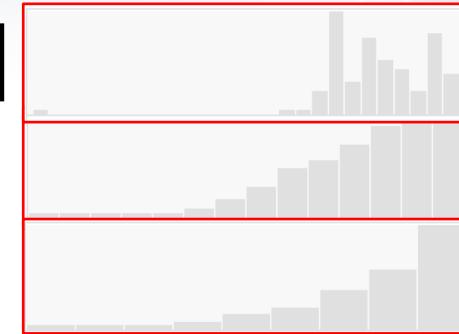
- **AR (absolute risk)** = the number of events (good or bad) in treated or control groups, divided by the number of people in that group
- **ARC** = the AR of events in the control group
- **ART** = the AR of events in the treatment group
- **ARR (absolute risk reduction)** = ARC – ART
- **RR (relative risk)** = ART / ARC
- **RRR (relative risk reduction)** = $(ARC - ART) / ARC$
- **RRR** = $1 - RR$
- **NNT (number needed to treat)** = $1 / ARR$



Terminology



- “Genetic risk profile” [71 hits]
- “Genetic risk prediction” [109 hits]
- “Genetic risk score” [962 hits]
- ”Polygenic risk score” [331 hits]



Genetic risk score is an estimate of the cumulative contribution of genetic factors to a specific outcome of interest in an individual that takes into account the reported risk alleles.

[Curr Protoc Hum Genet.](#) 2016 Oct
11;91:1.29.1-1.29.9. doi: 10.1002/cphg.20.



Traditional CVD Risk Factors



- age, sex, body mass index, blood lipid levels, blood pressure, and smoking status
- 10-year risk for developing cardiovascular disease



Framingham Risk Equation



NATIONAL CHOLESTEROL EDUCATION PROGRAM

Third Report of the Expert Panel on
Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III)



Risk Assessment Tool for Estimating Your 10-year Risk of Having a Heart Attack

The risk assessment tool below uses information from the Framingham Heart Study to predict a person's chance of having a heart attack in the next 10 years. This tool is designed for adults aged 20 and older who do not have heart disease or diabetes. To find your risk score, enter your information in the calculator below.

Age:

years

Female Male

Gender:

mg/dL

Total Cholesterol:

mg/dL

HDL Cholesterol:

No Yes

Smoker:

mm/Hg

Systolic Blood Pressure:

No Yes

Are you currently on any medication to treat high blood pressure.

Individual with diabetes is considered to have the same 10-year risk as individual with prior CAD

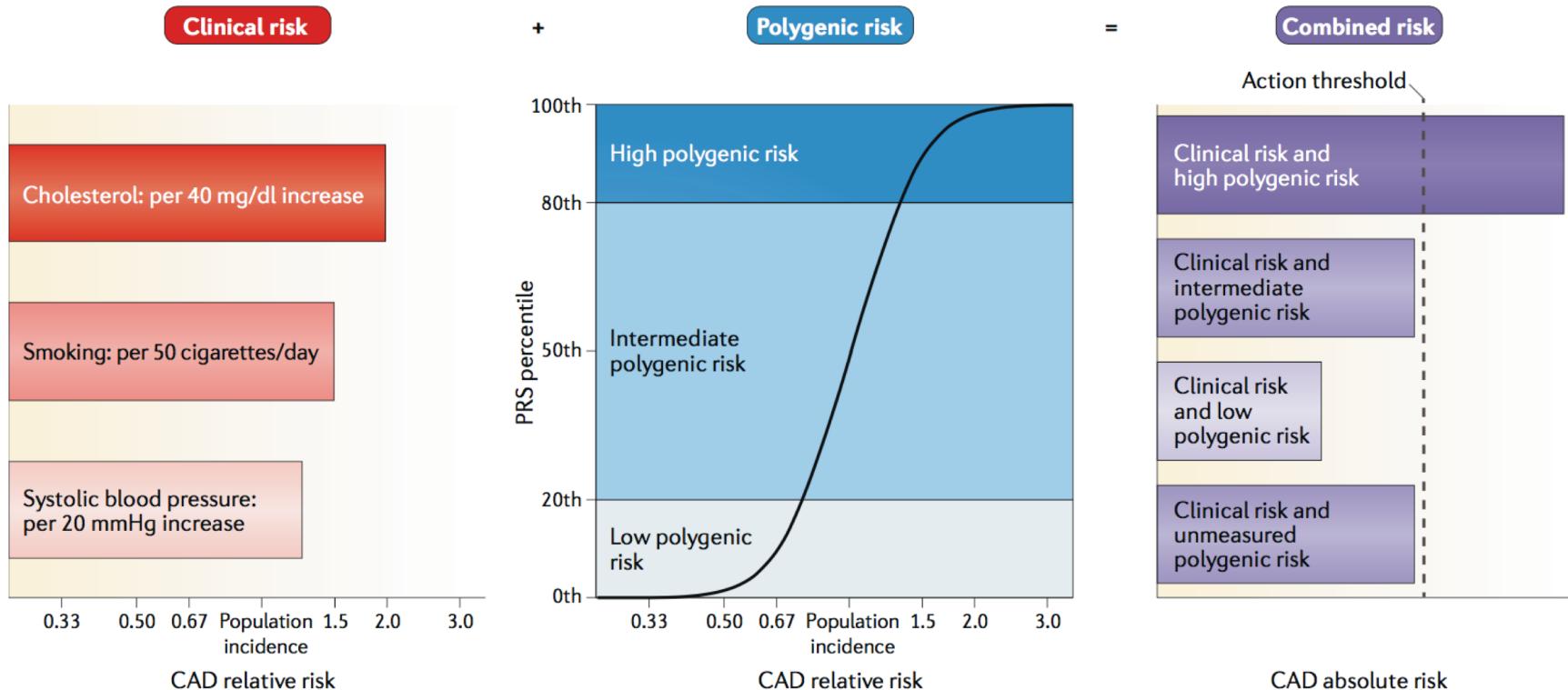


Rationale for PRS

- Early disease detection, prevention and intervention are fundamental goals for advancing human health.
- Knowledge of the genetic factors underlying these conditions has improved to a point where polygenic risk profiling on the basis of calculated polygenic risk scores (PRSs) provides personal and clinical utility.



Conventional + PRS



Clinical + PRS > risk : benefit ratio → Justify intervention

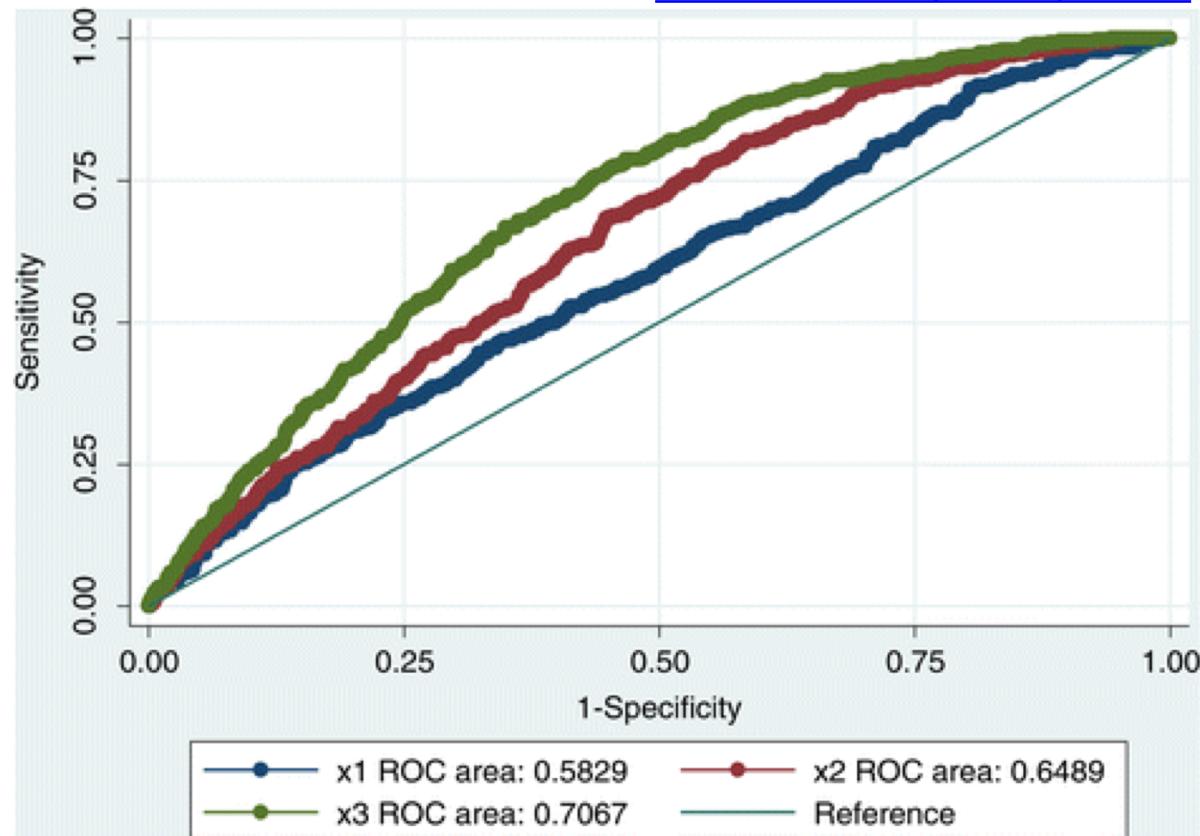


AUC -- Area Under the Curve

- Predictive ability of a receiver operating characteristic (ROC)

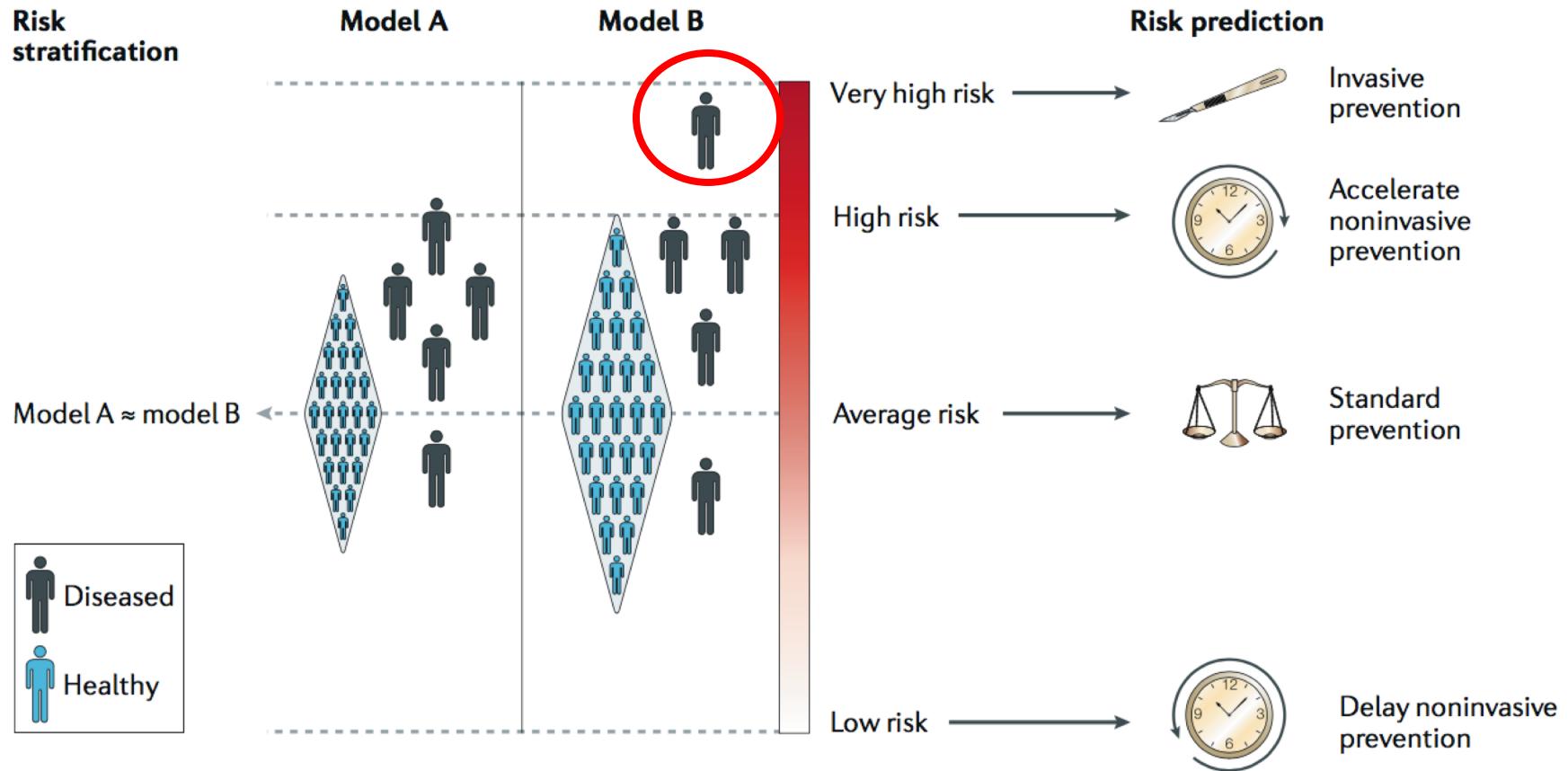
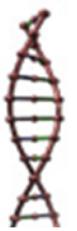
[BMC Medicine, 2015, 13:86](#)

ROC is the diagnostic ability of a binary classifier system as its discrimination threshold is varied.





Stratification vs Prediction



- Same AUC but different medical implication.



Utility of PRS

- Based on three classes of intervention
- PRS for therapeutic intervention
 - Help select appropriate intervention
- PRS to informed disease screening
 - PRS to help decide the time to start screening
- PRS for life planning
 - Promoting healthy behavior
 - Prediction of disease-onset might help with legal, financial, and care planning.





Building & Evaluating Absolute Risk Models



Discovery of risk factors

- 1 High-quality epidemiological studies with large sample sizes and refined and objective measurements of phenotypes and exposures are needed to identify novel risk factors (including genetic variation, environmental risk factors, biomarkers of exposure or internal dose).

Characterization of relative risk

- 2 Building of relative risk models that combine information on multiple risk factors (including polygenic risk scores, environmental risk factors and their interactions).

Estimation of absolute risk

- 3 Projecting risk of developing disease over a specified time interval based on a subject's risk factors (using relative risk models, distribution of risk factors, overall age-specific disease incidence and mortality rates in the target population).

Evaluation of model calibration

- 4 Comparison of the number of projected and observed disease diagnoses over a specified time period, within strata of people at different projected risk in prospective cohort studies.

Evaluation of public health utility

- 5 Evaluating effectiveness of primary and secondary prevention strategies tailored according to people's levels of projected risk.



A Common Allele on Chromosome 9 Associated with Coronary Heart Disease

Ruth McPherson,^{1,*†} Alexander Pertsemlidis,^{2,*} Nihan Kavaslar,¹ Alexandre Stewart,¹ Robert Roberts,¹ David R. Cox,³ David A. Hinds,³ Len A. Pennacchio,^{4,5} Anne Tybjaerg-Hansen,⁶ Aaron R. Folsom,⁷ Eric Boerwinkle,⁸ Helen H. Hobbs,^{2,9} Jonathan C. Cohen^{2,10†}

Coronary heart disease (CHD) is a major cause of death in Western countries. We used genome-wide association scanning to identify a 58-kilobase interval on chromosome 9p21 that was consistently associated with CHD in six independent samples (more than 23,000 participants) from four Caucasian populations. This interval, which is located near the *CDKN2A* and *CDKN2B* genes, contains no annotated genes and is not associated with established CHD risk factors such as plasma lipoproteins, hypertension, or diabetes. Homozygotes for the risk allele make up 20 to 25% of Caucasians and have a ~30 to 40% increased risk of CHD.

Mainly Canadian + US population

McPherson, et al. Science 316:1488-1491 (2007).



Screening

Genome-wide Association Scan (75,000 SNPs/person)

Ottawa Heart Study-1 (OHS-1)

322 Cases : 312 controls



Replicate Association Study 1: SNPs with $P < 0.025$

Ottawa Heart Study-2 (OHS-2)

311 cases : 326 controls



Replicate Association Study 2: SNPs with $P < 0.025$

Atherosclerosis Risk in Communities Study (ARIC)

1,347 cases : 9,054 controls

↓
rs10757274 and rs2383206

Validation

Copenhagen City Heart

Study (CCHS)

1,525 cases

9,053 controls

Dallas Heart Study

(DHS)

154 cases

527 controls

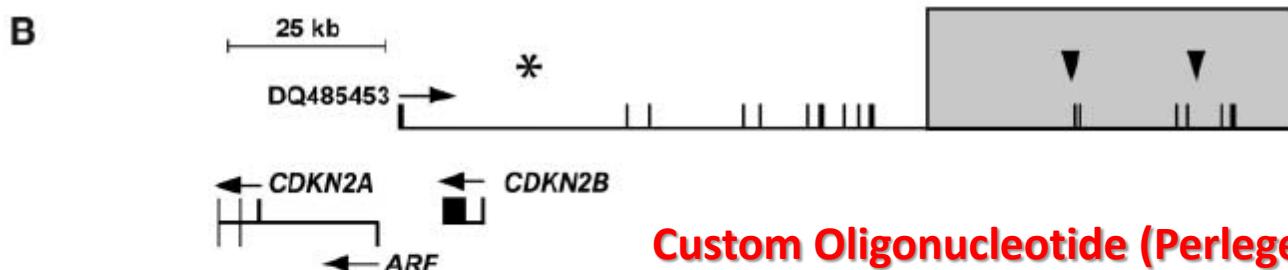
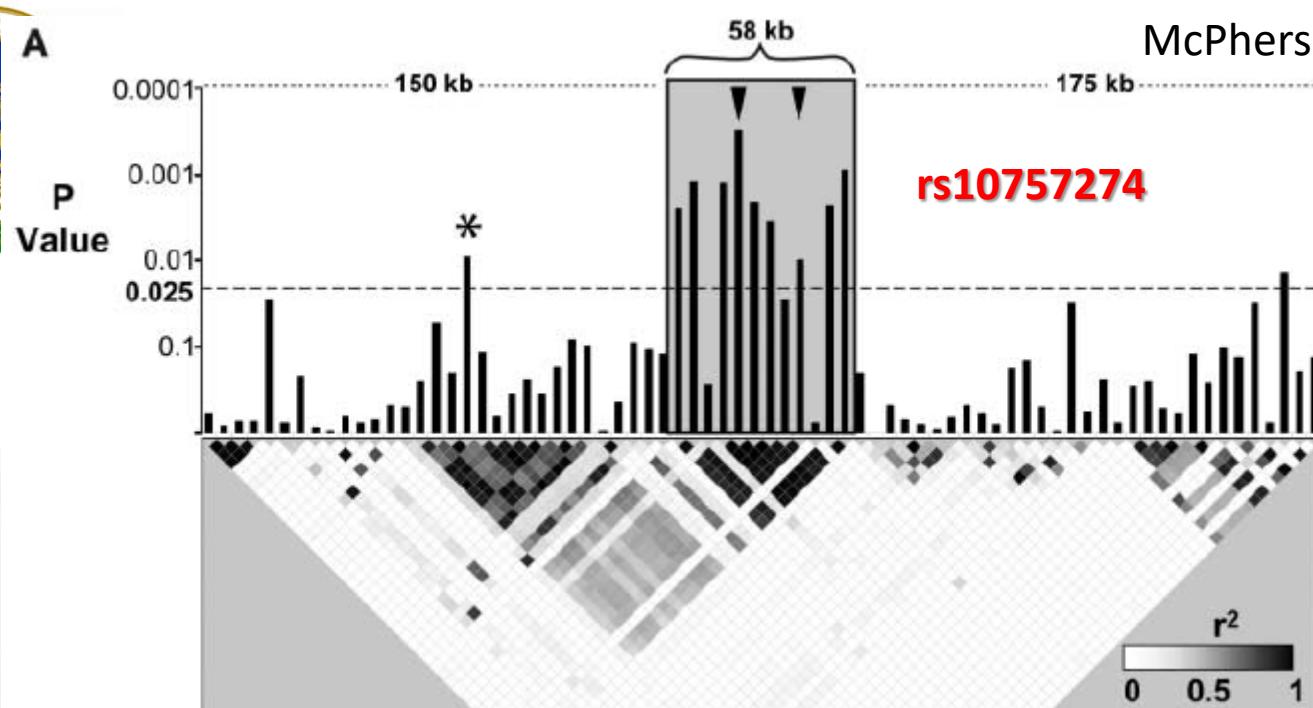
Ottawa Heart Study-3

(OHS-3)

647 cases

847 controls

Fig. 1. Study design for identification and validation of sequence variants associated with CHD. Assuming independence, the probability of any single SNP achieving a nominal significance level of 0.025 in all three studies with the associations being in the same direction was 3.9×10^{-6} ($0.025^3 \times 0.5^2$); thus, none of the 100,000 SNPs would be expected by chance to replicate consistently in all three comparisons.



Custom Oligonucleotide (Perlegen 100K)

~5 kb apart in the interval extending 175 kb upstream and downstream of rs10757274 and rs2383206 were assayed in 500 cases and 500 controls from the OHS population with GeneChip Human Mapping 500K Array Sets (Affymetrix, Santa Clara, CA). Bars represent P values (determined with χ^2 tests) for differences in allele frequency between cases and controls. Arrowheads indicate rs10757274 and rs2383206. The asterisk represents rs518394. The risk interval is indicated with a gray box. The linkage disequilibrium map indicates pairwise r^2 values. Blocks are shaded on a continuous scale, where white represents an r^2 of 0 and black represents an r^2 of 1. (B) Physical map of the region showing the location of the risk interval (gray box) relative to the noncoding RNA DQ485453 and adjacent genes CDKN2A, ARF, and CDKN2B. Arrowheads indicate rs10757274 and rs2383206, and the asterisk represents rs518394 [see (A)].





A Common Variant on Chromosome 9p21 Affects the Risk of Myocardial Infarction

Anna Helgadottir,^{1*} Gudmar Thorleifsson,^{1*} Andrei Manolescu,^{1*} Solveig Gretarsdottir,¹ Thorarinn Blöndal,¹ Aslaug Jonasdottir,¹ Adalbjorg Jonasdottir,¹ Asgeir Sigurdsson,¹ Adam Baker,¹ Arnar Palsson,¹ Gisli Masson,¹ Daniel F. Gudbjartsson,¹ Kristinn P. Magnusson,¹ Karl Andersen,² Allan I. Levey,³ Valgerdur M. Backman,¹ Sigurborg Matthiasdottir,¹ Thorbjorg Jónsdóttir,¹ Stefan Palsson,¹ Helga Einarsdottir,¹ Steinunn Gunnarsdottir,¹ Arnaldur Gylfason,¹ Viola Vaccarino,³ W. Craig Hooper,³ Muredach P. Reilly,⁴ Christopher B. Granger,⁵ Harland Austin,³ Daniel J. Rader,⁴ Svali H. Shah,⁵ Arshed A. Quyyumi,³ Jeffrey R. Gulcher,¹ Gudmundur Thorgeirsson,² Unnur Thorsteinsdottir,¹ Augustine Kong,^{1†} Kari Stefansson^{1†}

The global endemic of cardiovascular diseases calls for improved risk assessment and treatment. Here, we describe an association between myocardial infarction (MI) and a common sequence variant on chromosome 9p21. This study included a total of 4587 cases and 12,767 controls. The identified variant, adjacent to the tumor suppressor genes *CDKN2A* and *CDKN2B*, was associated with the disease with high significance. Approximately 21% of individuals in the population are homozygous for this variant, and their estimated risk of suffering myocardial infarction is 1.64 times as great as that of noncarriers. The corresponding risk is 2.02 times as great for early-onset cases. The population attributable risk is 21% for MI in general and 31% for early-onset cases.

Helgadottir, et al. Science 316:1491-1493 (2007).



rs10757278 and CAD



Table 2. Genotype-specific OR for the risk allele of rs10757278. The risk for heterozygous carriers (0X) and homozygous carriers (XX) is compared with the risk for noncarriers (00), together with 95% CI and the population attributable risk (PAR). The lower part of the table includes the corresponding values when the analysis is restricted to early-onset MI cases. Study population includes the number of MI cases (*n*) and controls (*m*). For the Icelandic groups, *P* values and OR were adjusted for relatedness using simulations.

Study population (<i>n/m</i>) Variant (allele)	Genotype-specific OR			PAR
	00	0X (95% CI)	XX (95% CI)	
Iceland (2272/10,261) rs10757278 (G)	1	1.25 (1.12–1.39)	1.58 (1.38–1.81)	0.19
U.S. groups (2315/2506) rs10757278 (G)	1	1.28 (1.14–1.45)	1.72 (1.45–2.03)	0.23
All groups (4587/12,767) rs10757278 (G)	1	1.26 (1.16–1.36)	1.64 (1.47–1.82)	0.21
<i>Early-onset MI (<50 for males; <60 for females)</i>				
Iceland (621/10,261) rs10757278 (G)	1	1.38 (1.13–1.69)	1.94 (1.53–2.46)	0.27
U.S. groups (1080/2506) rs10757278 (G)	1	1.56 (1.32–1.85)	2.08 (1.69–2.58)	0.34
All groups (1701/12,767) rs10757278 (G)	1	1.49 (1.31–1.69)	2.02 (1.72–2.36)	0.31



A

The NEW ENGLAND JOURNAL *of* MEDICINE

ESTABLISHED IN 1812

AUGUST 2, 2007

VOL. 357 NO. 5

Genomewide Association Analysis of Coronary Artery Disease

Nilesh J. Samani, F.Med.Sci., Jeanette Erdmann, Ph.D., Alistair S. Hall, F.R.C.P., Christian Hengstenberg, M.D., Massimo Mangino, Ph.D., Bjoern Mayer, M.D., Richard J. Dixon, Ph.D., Thomas Meitinger, M.D., Peter Braund, M.Sc., H.-Erich Wichmann, M.D., Jennifer H. Barrett, Ph.D., Inke R. König, Ph.D., Suzanne E. Stevens, M.Sc., Silke Szymczak, M.Sc., David-Alexandre Tregouet, Ph.D., Mark M. Iles, Ph.D., Friedrich Pahlke, M.Sc., Helen Pollard, M.Sc., Wolfgang Lieb, M.D., Francois Cambien, M.D., Marcus Fischer, M.D., Willem Ouwehand, F.R.C.Path., Stefan Blankenberg, M.D., Anthony J. Balmforth, Ph.D., Andrea Baessler, M.D., Stephen G. Ball, F.R.C.P., Tim M. Strom, M.D., Ingrid Brænne, M.Sc., Christian Gieger, Ph.D., Panos Deloukas, Ph.D., Martin D. Tobin, M.F.P.H.M., Andreas Ziegler, Ph.D., John R. Thompson, Ph.D., and Heribert Schunkert, M.D., for the WTCCC and the Cardiogenics Consortium*

Samani, et al. N Engl J Med 357(5): 443-53 (2007).



rs1333049 and CAD



BACKGROUND

Modern genotyping platforms permit a systematic search for inherited components of complex diseases. We performed a joint analysis of two genomewide association studies of coronary artery disease.

METHODS

We first identified chromosomal loci that were strongly associated with coronary artery disease in the Wellcome Trust Case Control Consortium (WTCCC) study (which involved 1926 case subjects with coronary artery disease and 2938 controls) and looked for replication in the German MI [Myocardial Infarction] Family Study (which involved 875 case subjects with myocardial infarction and 1644 controls). Data on other single-nucleotide polymorphisms (SNPs) that were significantly associated with coronary artery disease in either study ($P<0.001$) were then combined to identify additional loci with a high probability of true association. Genotyping in both studies was performed with the use of the GeneChip Human Mapping 500K Array Set (Affymetrix).

RESULTS

Of thousands of chromosomal loci studied, the same locus had the strongest association with coronary artery disease in both the WTCCC and the German studies: chromosome 9p21.3 (SNP, rs1333049) ($P=1.80\times10^{-14}$ and $P=3.40\times10^{-6}$, respectively). Overall, the WTCCC study revealed nine loci that were strongly associated with coronary artery disease ($P<1.2\times10^{-5}$ and less than a 50% chance of being falsely positive). In addition to chromosome 9p21.3, two of these loci were successfully replicated (adjusted $P<0.05$) in the German study: chromosome 6q25.1 (rs6922269) and chromosome 2q36.3 (rs2943634). The combined analysis of the two studies identified four additional loci significantly associated with coronary artery disease ($P<1.3\times10^{-6}$) and a high probability (>80%) of a true association: chromosomes 1p13.3 (rs599839), 1q41 (rs17465637), 10q11.21 (rs501120), and 15q22.33 (rs17228212).

CONCLUSIONS

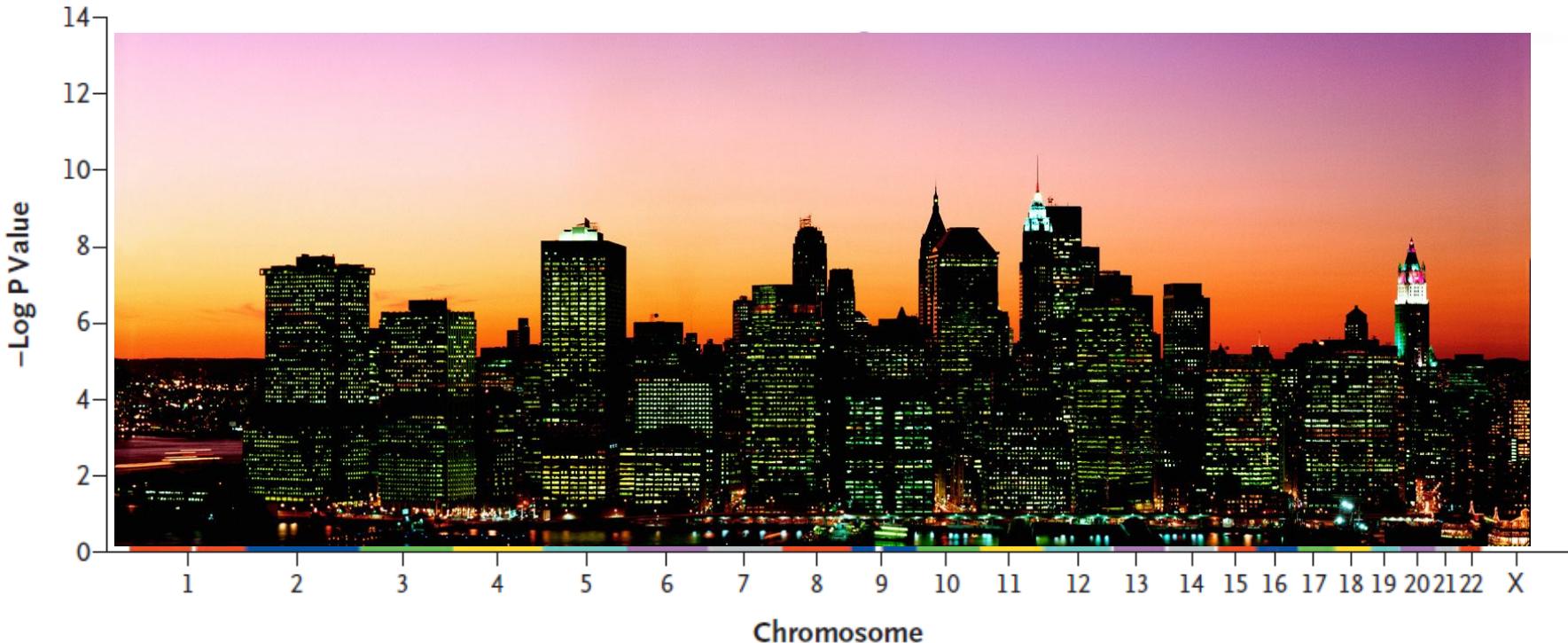
We identified several genetic loci that, individually and in aggregate, substantially affect the risk of development of coronary artery disease.



GWAS Signal Plot

P plot: height = $-\log(p\text{-value})$
= Strength of the signals

A WTCCC Study



Manhattan skyline?



Linkage disequilibrium structure within the region of 9p21.3



- Genotyped SNPs
- Entrez genes

NM_000077
CDKN2A: cyclin-dependent kinase inhibitor 2A isoform 1
NM_058195
CDKN2A: alternative reading frame p14 isoform 4
NM_058197
CDKN2A: cyclin-dependent kinase inhibitor 2A isoform 3
NM_004936
CDKN2B: cyclin-dependent kinase inhibitor
NM_078487
CDKN2B: cyclin-dependent kinase inhibitor 2B isoform 2

rs10757274@22086055*

rs10757278@22114477deCODE

rs1333049@22115503 WTCC

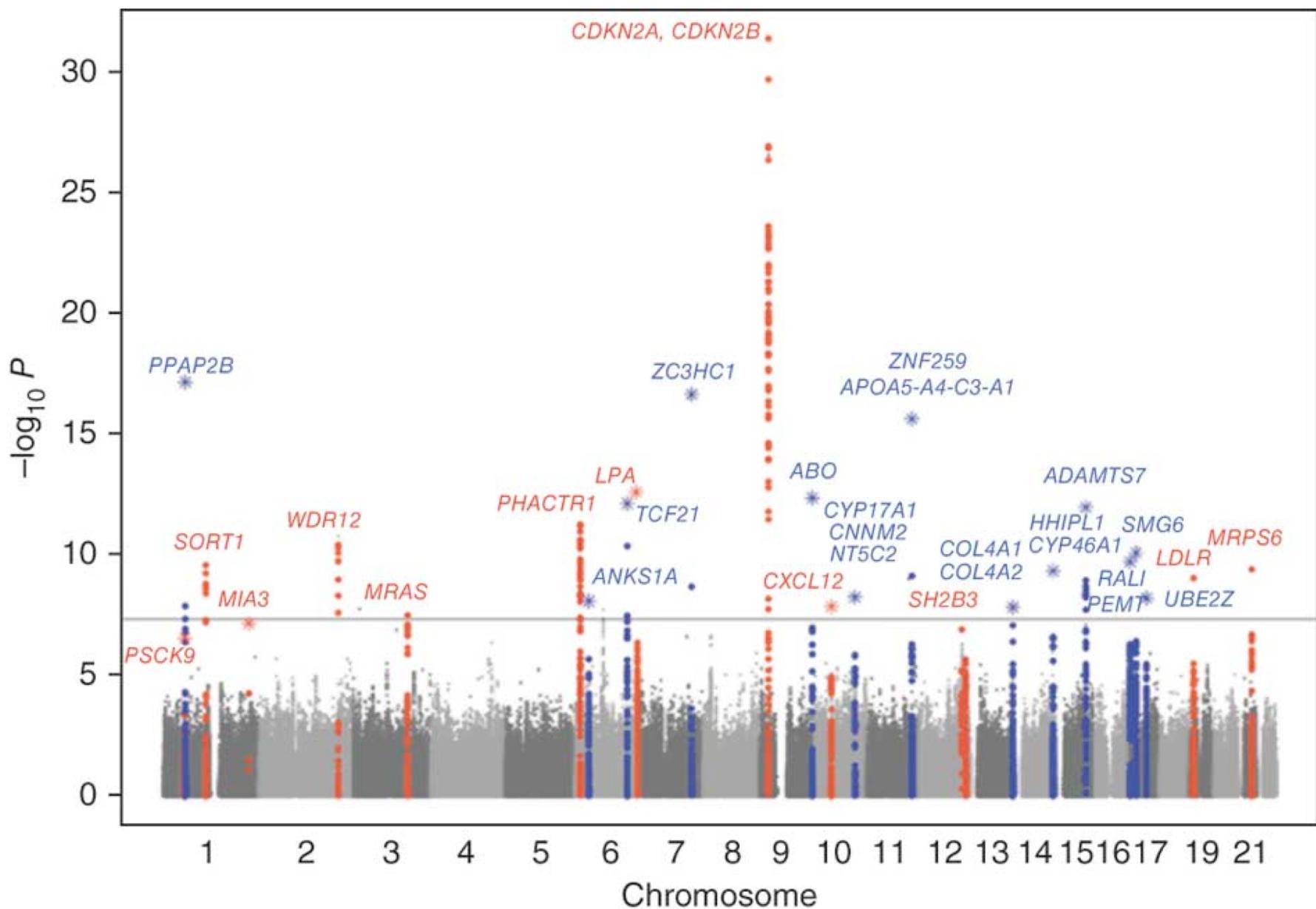
- LD Plot

CEU: lod

Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease

Heribert Schunkert, Inke R König, Sekar Kathiresan, Muredach P Reilly, Themistocles L Assimes, Hilma Holm, Michael Preuss, Alexandre F R Stewart, Maja Barbalic, Christian Gieger, Devin Absher, Zouhair Aherrahrou, Hooman Allayee, David Altshuler, Sonia S Anand, Karl Andersen, Jeffrey L Anderson, Diego Ardissino, Stephen G Ball, Anthony J Balmforth, Timothy A Barnes, Diane M Becker, Lewis C Becker, Klaus Berger, Joshua C Bis ... *for the CARDIoGRAM Consortium*

We performed a meta-analysis of 14 genome-wide association studies of coronary artery disease (CAD) comprising 22,233 individuals with CAD (cases) and 64,762 controls of European descent followed by genotyping of top association signals in 56,682 additional individuals. This analysis identified 13 loci newly associated with CAD at $P < 5 \times 10^{-8}$ and confirmed the association of 10 of 12 previously reported CAD loci. The 13 new loci showed risk allele frequencies ranging from 0.13 to 0.91 and were associated with a 6% to 17% increase in the risk of CAD per allele. Notably, only three of the new loci showed significant association with traditional CAD risk factors and the majority lie in gene regions not previously implicated in the pathogenesis of CAD. Finally, five of the new CAD risk loci appear to have pleiotropic effects, showing strong association with various other human diseases or traits.



Known loci are shown in red and newly discovered loci are shown in blue.

Large-scale association analysis identifies new risk loci for coronary artery disease

The CARDIoGRAMplusC4D Consortium, Panos Deloukas  [...] Nilesh J Samani 

Nature Genetics 45, 25–33 (2013) | Download Citation  

Coronary artery disease (CAD) is the commonest cause of death. Here, we report an association analysis in **63,746 CAD cases and 130,681 controls** identifying **15 loci** reaching genome-wide significance, taking **the number of susceptibility loci for CAD to 46**, and a further 104 independent variants ($r^2 < 0.2$) strongly associated with CAD at a 5% false discovery rate (FDR). Together, these variants explain approximately 10.6% of CAD heritability. Of the 46 genome-wide significant lead SNPs, 12 show a significant association with a lipid trait, and 5 show a significant association with blood pressure, but none is significantly associated with diabetes. Network analysis with 233 candidate genes (loci at 10% FDR) generated 5 interaction networks comprising 85% of these putative genes involved in CAD. The four most significant pathways mapping to these networks are linked to lipid metabolism and inflammation, underscoring the causal role of these activities in the genetic etiology of CAD. Our study provides insights into the genetic basis of CAD and identifies key biological pathways.

A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease

the CARDIoGRAMplusC4D Consortium

Existing knowledge of genetic variants affecting risk of coronary artery disease (CAD) is largely based on genome-wide association study (GWAS) analysis of common SNPs. Leveraging phased haplotypes from the 1000 Genomes Project, we report a GWAS meta-analysis of ~185,000 CAD cases and controls, interrogating 6.7 million common (minor allele frequency (MAF) > 0.05) and 2.7 million low-frequency (0.005 < MAF < 0.05) variants. In addition to confirming most known CAD-associated loci, we identified ten new loci (eight additive and two recessive) that contain candidate causal genes newly implicating biological processes in vessel walls. We observed intralocus allelic heterogeneity but little evidence of low-frequency variants with larger effects and no evidence of synthetic association. Our analysis provides a comprehensive survey of the fine genetic architecture of CAD, showing that genetic susceptibility to this common disease is largely determined by common SNPs of small effect size.



Polygenic Risk Score

Polygenic Risk Score Identifies Subgroup With Higher Burden of Atherosclerosis and Greater Relative Benefit From Statin Therapy in the Primary Prevention Setting

Pradeep Natarajan, Robin Young, Nathan O. Stitziel,
Sandosh Padmanabhan, Usman Baber, Roxana Mehran, Samantha Sartori,
Valentin Fuster, Dermot F. Reilly, Adam Butterworth, Daniel J. Rader,
Ian Ford, Naveed Sattar, and Sekar Kathiresan Show less Authors

Originally published 21 Feb 2017 | Circulation. 2017;135:2091–2101



Polygenic Risk Score Development



- 67 SNPs associated with CHD from GWAS
- **Fifty-seven** of these variants were **genotyped** among WOSCOPS participants with the Illumina Metabochip¹⁸
- **Thirteen variants** were directly **genotyped** among CARDIA participants with the Affymetrix Human SNP Array 6.0 and another **25 proxy variants** available through **statistical imputation**.
- **Fifty-nine variants** were directly genotyped among Biolmage participants with the Illumina HumanExome Beadchip¹⁹ and an additional **4 proxy variants** ($r^2 > 0.8$) available through **statistical imputation**.



Where did these 67 SNPs come from?



- Nature Genetics 45(1), 25–33 (2013)
doi:10.1038/ng.2480
- Nature Genetics 47, 1121–1130 (2015)
doi:10.1038/ng.3396
- N Engl J Med. 2016 Mar 24;374(12):1134-44. doi: 10.1056/NEJMoa1507652.



67 Variants associated with CAD

	Locus	Implicated Gene(s)	WOSCOPS			BioImage			CARDIA		
			Lead SNP	Lead Risk Allele	CHD OR	SNP	Risk Allele	Frequency	SNP	Risk Allele	Frequency
1	1p13.3	<i>SORT1</i>	rs602633	C	1.12	rs602633	C	0.793	rs602633	C	0.793
..	1p32.2	<i>PPAP2B</i>	rs17114036	A	1.11	rs17114036	A	0.879	rs17114036	A	0.879
.	1p32.3	<i>PCSK9</i>	rs11206510	T	1.06	rs11206510	T	0.828	rs11206510	T	0.828
.	1q21.3	<i>IL6R</i>	rs4845625	T	1.06	rs4845625	T	0.451	rs4845625	T	0.451
.	1q41	<i>MIA3</i>	rs17464857	T	1.05	rs17464857	T	0.866	rs17464857	T	0.866
.	21q22.11	<i>KCNE2</i>	rs9982601	T	1.13	rs9982601	T	0.122	rs9982601	T	0.122
.	22q11.23	<i>POM121L9P-ADORA2B</i>	rs180803	G	1.02						
.	2p11.2	<i>VAMP5-VAMP8-GGCX</i>	rs1561198	A	1.06	rs1561198	A	0.476	rs1561198	A	0.476
.	2p21	<i>ABCG5-ABCG8</i>	rs6544713	T	1.06	rs6544713	T	0.309	rs6544713	T	0.309
.	2p24.1	<i>APOB</i>	rs515135	G	1.07	rs515135	G	0.796	rs515135	G	0.796
.	2q22.3	<i>ZEB2-AC074093.1</i>	rs2252641	G	1.06	rs2252641	G	0.46	rs2252641	G	0.46
.	2q33.2	<i>WDR12</i>	rs6725887	C	1.12	rs6725887	C	0.13	rs2351524	T	0.13
.	2q37.1	<i>KCNJ13-GIGYF2</i>	rs1801251	A	1.05				rs1801251	A	0.287
.	3q22.3	<i>MRAS</i>	rs9818870	T	1.07	rs9818870	T	0.143	rs9818870	T	0.143
.	4q12	<i>REST-NOAI</i>	rs17087335	T	1.06	rs17087335	T	0.186	rs17087335	T	0.186
.	4q31.22	<i>EDNRA</i>	rs1878406	T	1.1	rs1878406	T	0.146			
.	4q32.1	<i>GUCY1A3</i>	rs7692387	G	1.08	rs7692387	G	0.795	rs7692387	G	0.795
.	5q31.1	<i>SLC22A4-SLC22A5</i>	rs273909	C	1.07	rs273909	C	0.134	rs273909	C	0.134
.	6p21.2	<i>KCNK5</i>	rs10947789	T	1.07	rs10947789	T	0.736	rs10947789	T	0.736
.	6p21.31	<i>ANKS1A</i>	rs12205331	C	1.04	rs12205331	C	0.81	rs12205331	C	0.81
.	6p21.33	<i>C2</i>	rs3130683	T	1.09				rs3130683	T	0.968
.	6p24.1	<i>PHACTR1</i>	rs9369640	A	1.09	rs9369640	A	0.62	rs9369640	A	0.62
.	6q23.2	<i>TCF21</i>	rs12190287	C	1.07	rs12190287	C	0.612			
.	6q25.3	<i>LPA</i>	rs3798220	C	1.28	rs3798220	C	0.01	rs3798220	C	0.01
.	6q25.3	<i>LPA</i>	rs2048327	G	1.06	rs2048327	G	0.352	rs2048327	G	0.352



67 Variants associated with CAD



27	6q26	<i>PLG</i>	rs4252120	T	1.07	rs4252120	T	0.704	rs4252120	T	0.704	rs4252120	T	0.704
..	7p21.1	<i>HDAC9</i>	rs2023938	G	1.08	rs2023938	G	0.095	rs2023938	G	0.095	<i>rs10245779</i>	C	0.1
.	7q22.3	<i>NA</i>	rs12539895	A	1.08	rs12539895	A	0.188	<i>rs3815148</i>	C	0.203	<i>rs34084719</i>	C	0.188
.	7q32.2	<i>ZC3HC1</i>	rs11556924	C	1.09	rs11556924	C	0.623	rs11556924	C	0.623			
.	7q36.1	<i>NOS3</i>	rs3918226	T	1.14	rs3918226	T	0.014						
.	8p21.3	<i>LPL</i>	rs264	G	1.11	rs264	G	0.851	rs264	G	0.851	<i>rs271</i>	G	0.847
.	8q24.13	<i>TRIB1</i>	rs2954029	A	1.06	rs2954029	A	0.552	rs2954029	A	0.552	<i>rs2980875</i>	A	0.548
..	9p21.3	<i>CDKN2A</i>	rs1333049	C	1.21	rs1333049	C	0.472	rs1333049	C	0.472	rs1333049	C	0.472
.	9p21.3	<i>CDKN2A</i>	rs3217992	A	1.16	rs3217992	A	0.402	rs3217992	A	0.402	rs3217992	A	0.402
.	9q31.3	<i>SVEPI</i>	rs111245230	C	1.14				rs111245230	C	0.032	<i>rs11791314</i>	C	0.029
.	9q34.2	<i>ABO</i>	rs579459	C	1.07	rs579459	C	0.214	rs579459	C	0.214			
.	10p11.23	<i>KIAA1462</i>	rs2505083	C	1.06	rs2505083	C	0.428	rs2505083	C	0.428			
..	10q11.21	<i>CXCL12</i>	rs501120	A	1.07	rs501120	A	0.854	rs501120	A	0.854	<i>rs671765</i>	A	0.854
.	10q11.21	<i>CXCL12</i>	rs2047009	C	1.05	rs2047009	C	0.49				rs2047009	C	0.49
.	10q23.31	<i>LIPA</i>	rs2246833	T	1.06	rs2246833	T	0.358	<i>rs2246942</i>	G	0.361			
.	10q23.31	<i>LIPA</i>	rs11203042	T	1.04	rs11203042	T	0.457	rs11203042	T	0.457	<i>rs11203041</i>	T	0.455
.	10q24.32	<i>CYP17A1</i>	rs12413409	G	1.1	rs12413409	G	0.911	rs12413409	G	0.911	<i>rs12052058</i>	G	0.746
.	11p15.3	<i>NA</i>	rs11042937	T	1.04				rs11042937	T	0.518	rs11042937	T	0.518
.	11p15.4	<i>SWAP70</i>	rs10840293	A	1.06									
.	11q22.3	<i>PDGF</i>	rs974819	A	1.07	rs974819	A	0.279	rs974819	A	0.279	<i>rs2128739</i>	A	0.272
.	11q23.3	<i>APOA5-APOA1</i>	rs9326246	C	1.09	rs9326246	C	0.093	rs9326246	C	0.093	<i>rs6589566</i>	G	0.091
.	12p13.3	<i>LRPI</i>	rs11172113	C	1.06	rs11172113	C	0.388	rs11172113	C	0.388			
.	12p24.31	<i>SCARB1</i>	rs11057830	A	1.08	rs11057830	A	0.165	rs11057830	A	0.165			
.	12q24.12	<i>SH2B3</i>	rs3184504	T	1.07	rs3184504	T	0.464	rs3184504	T	0.464			
.	13q12.3	<i>FLT1</i>	rs9319428	A	1.06	rs9319428	A	0.309	rs9319428	A	0.309			
.	13q34	<i>COL4A1</i>	rs9515203	T	1.08	rs9515203	T	0.766						
.	13q34	<i>COL4A1</i>	rs4773144	G	1.07	rs4773144	G	0.447	rs4773144	G	0.447			
.	14q32.2	<i>HHIP1</i>	rs2895811	C	1.06	rs2895811	C	0.426	rs2895811	C	0.426			
.	15q22.33	<i>SMAD3</i>	rs17293632	C	1.05				rs17293632	C	0.785	<i>rs2033784</i>	T	0.719
56	15q25.1	<i>ADAMTS7</i>	rs7173743	T	1.07	rs7173743	T	0.535	rs7173743	T	0.535	<i>rs7168915</i>	A	0.533



67 Variants associated with CAD

57	15q26.1	<i>FURIN</i>	rs17514846	A	1.07	rs17514846	A	0.461	rs17514846	A	0.461	<i>rs1894401</i>	<i>G</i>	0.469
58	16q13	<i>CETP</i>	rs247616	C	1.05	rs247616	C	0.708	rs247616	C	0.708			
59	17p11.2	<i>RASDI</i>	rs12936587	G	1.06	rs12936587	G	0.554	rs12936587	G	0.554			
60	17p13.3	<i>SMG6</i>	rs2281727	C	1.04	rs2281727	C	0.369	rs2281727	C	0.369	<i>rs9895551</i>	<i>G</i>	0.37
61	17q21.32	<i>UBE2Z</i>	rs15563	C	1.04	rs15563	C	0.559	rs15563	C	0.559	rs15563	C	0.559
62	17q23.2	<i>BCAS3</i>	rs8080784	C	1.06				rs8080784	C	0.161	<i>rs7225581</i>	<i>A</i>	0.161
63	18q21.32	<i>PMAIP1-MC4R</i>	rs663129	A	1.06	rs663129	A	0.24	<i>rs571312</i>	<i>T</i>	0.24			
64	19p13.2	<i>ANGPTL4</i>	rs116843064	G	1.16				rs116843064	G	0.974			
65	19p13.2	<i>LDLR</i>	rs1122608	G	1.1	rs1122608	G	0.75	rs1122608	G	0.75	rs1122608	G	0.75
66	19q13.32	<i>APOE-APOC1</i>	rs445925	C	1.13	rs445925	C	0.906	rs445925	C	0.906			
67	19q13.32	<i>APOE-APOC1</i>	rs2075650	G	1.11	rs2075650	G	0.131	rs2075650	G	0.131			

A set of 67 independent genomic variants associated with coronary heart disease at stringent statistical thresholds have been found.¹⁻³ Directly genotyped variants in WOSCOPS were used for analysis. Directly genotyped and imputed proxy variants ($r^2 > 0.8$)⁴ in BioImage and CARDIA were included for analysis. Variants in red and italicized are proxy variants. Risk allele frequencies in 1000G phase 3 EUR samples are noted.⁴



Polygenic Risk Score Construction



A polygenic risk score was constructed by weighting the total number of risk alleles by their effects (log of the odds ratios) of CHD risk from the published literature. Incremental scores from missing genotypes in individuals were imputed from the allele frequency in each cohort. To account for the differences in the numbers of variants per cohort, a normalized polygenic risk score (mean=0, SD=1) was created per cohort (**Figure I in the online-only Data Supplement**).



Genetic Risk Score

$$OR = e^{\sum_{i=1}^n \beta_i X_i}$$

	ATS-1	ATS-2	ATS-3	ATS-4	ATS-5
CAD Risk	2.94	2.77	2.96	2.44	2.54

Your Genetic Risk,
Compared to a person with no risk variants



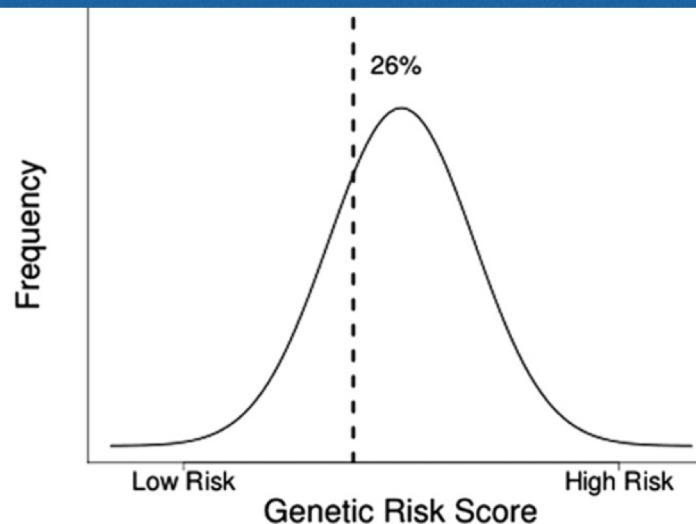
Your Risk Score

Based on the traditional Framingham risk score, your risk of coronary heart disease over the next 10 years is approximately 9.7%.

We tested for a total of 38 possible risk variants or alleles. Out of these 38, you carry 15 variants that are associated with higher risk.

Your genetic profile puts you in the 26 percentile for risk. This means 26% of the general population have a genetic risk score more favorable than you and 74% have a genetic risk score less favorable than you.

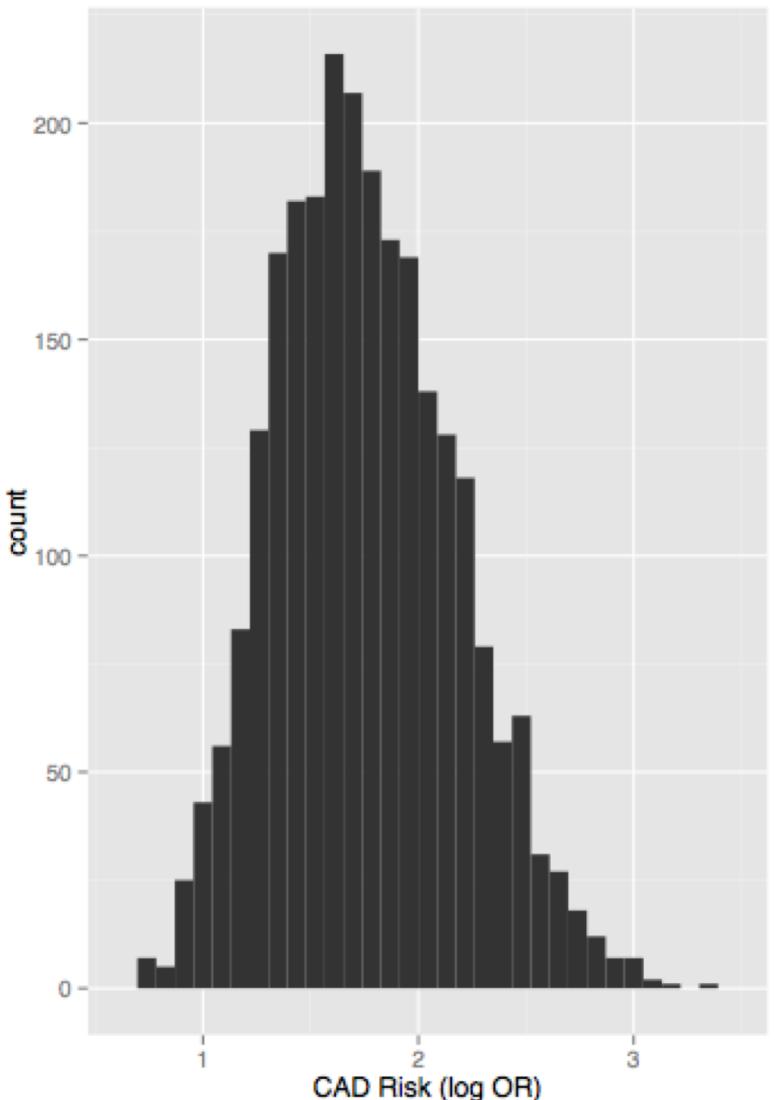
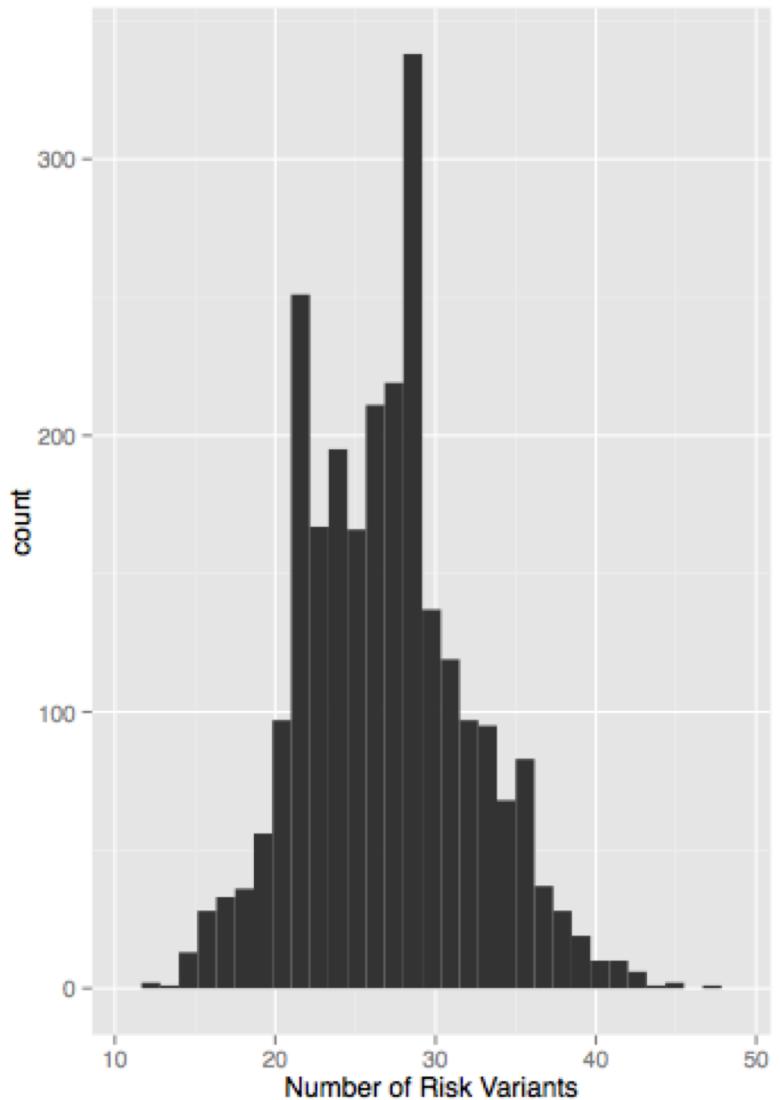
Your Genetic Risk Compared to the General Population

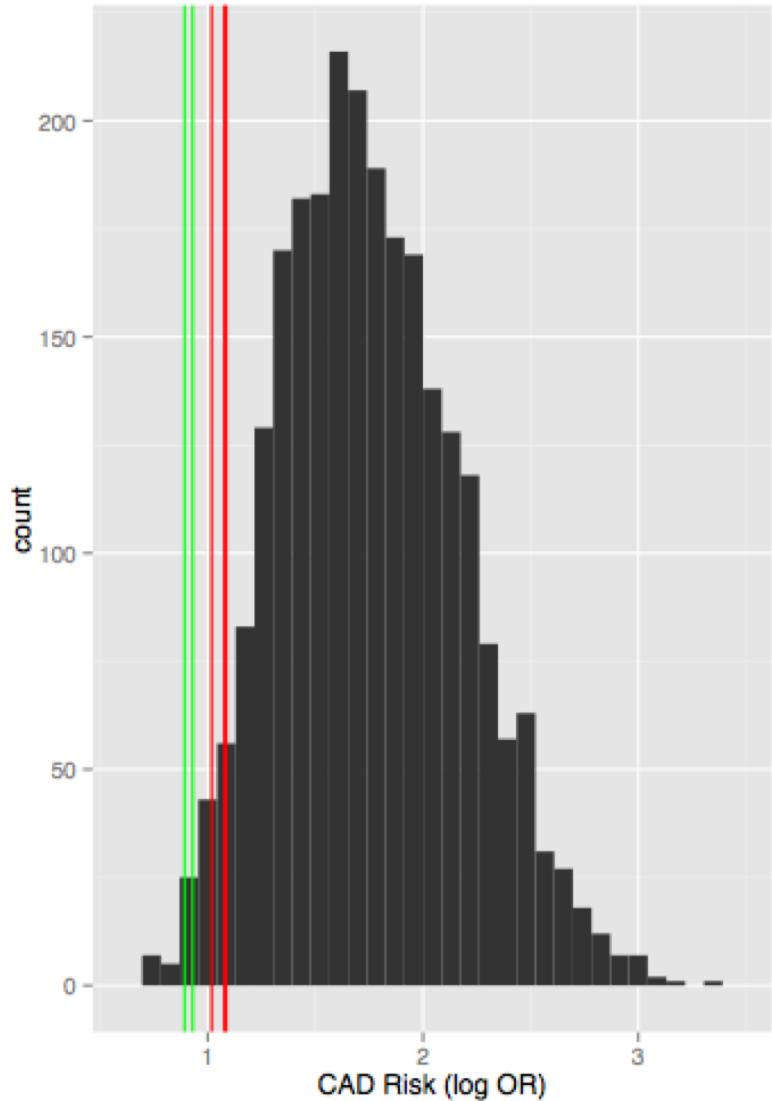


Based on the traditional Framingham risk score plus the genetic risk score, your risk of coronary heart disease over the next 10 years is approximately 8.7%.



CAD Risk Variants & CAD Risk Distribution in 1000 Genome Project





ATS-1	4.0
ATS-2	2.6
ATS-3	4.2
ATS-4	0.7
ATS-5	1.1





Building & Evaluating Absolute Risk Models



Discovery of risk factors

- 1 High-quality epidemiological studies with large sample sizes and refined and objective measurements of phenotypes and exposures are needed to identify novel risk factors (including genetic variation, environmental risk factors, biomarkers of exposure or internal dose).

Characterization of relative risk

- 2 Building of relative risk models that combine information on multiple risk factors (including polygenic risk scores, environmental risk factors and their interactions).

Estimation of absolute risk

- 3 Projecting risk of developing disease over a specified time interval based on a subject's risk factors (using relative risk models, distribution of risk factors, overall age-specific disease incidence and mortality rates in the target population).

Evaluation of model calibration

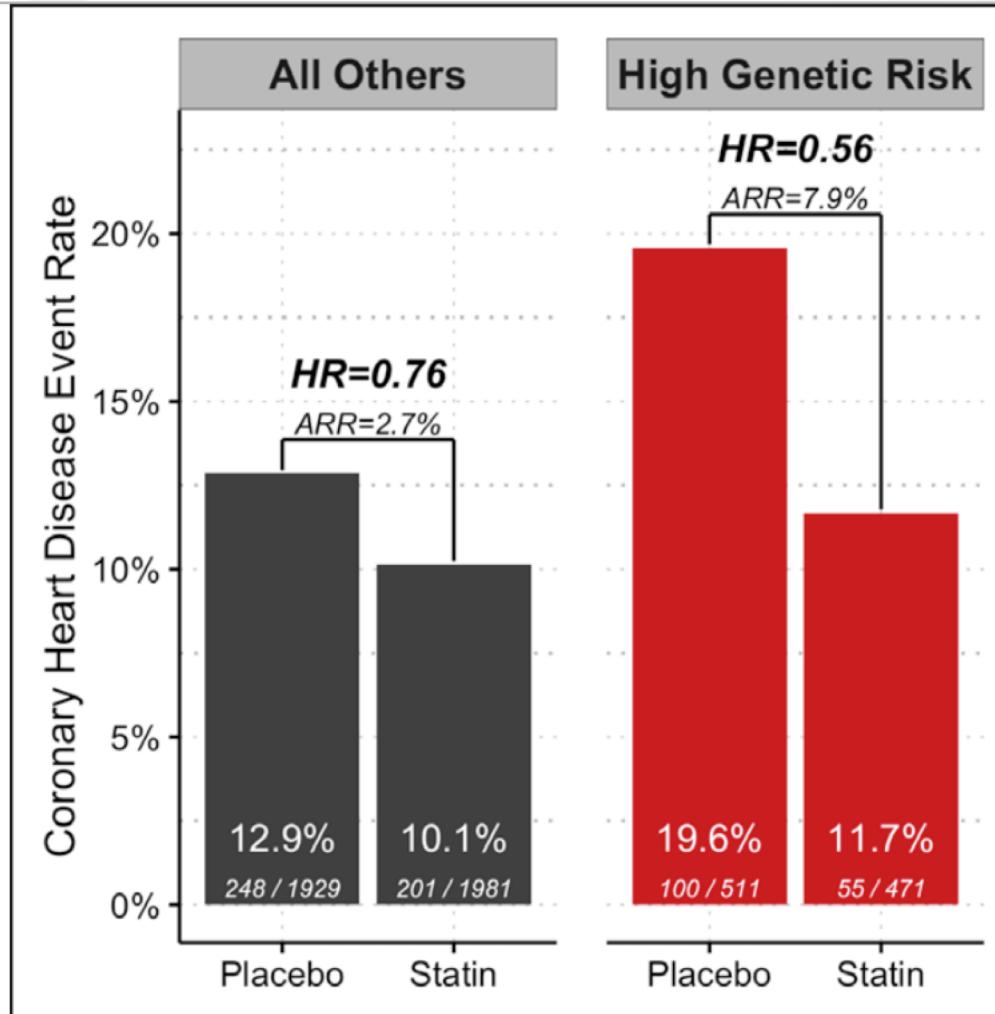
- 4 Comparison of the number of projected and observed disease diagnoses over a specified time period, within strata of people at different projected risk in prospective cohort studies.

Evaluation of public health utility

- 5 Evaluating effectiveness of primary and secondary prevention strategies tailored according to people's levels of projected risk.



Absolute Risk Estimation



Incident coronary heart disease events by statin therapy and genetic risk group in WOSCOPS (West of Scotland Coronary Prevention Study).

Genomic prediction of coronary heart disease

Gad Abraham^{1,2}, Aki S. Havulinna³, Oneil G. Bhalala^{1,2}, Sean G. Byars^{1,2}, Alysha M. De Livera^{1,2,4}, Laxman Yetukuri⁵, Emmi Tikkanen⁵, Markus Perola^{3,5}, Heribert Schunkert^{6,7}, Eric J. Sijbrands⁸, Aarno Palotie^{5,9,10,11}, Nilesh J. Samani^{12,13*}, Veikko Salomaa^{3*}, Samuli Ripatti^{5,14,15*†} and Michael Inouye^{1,2,5*†}

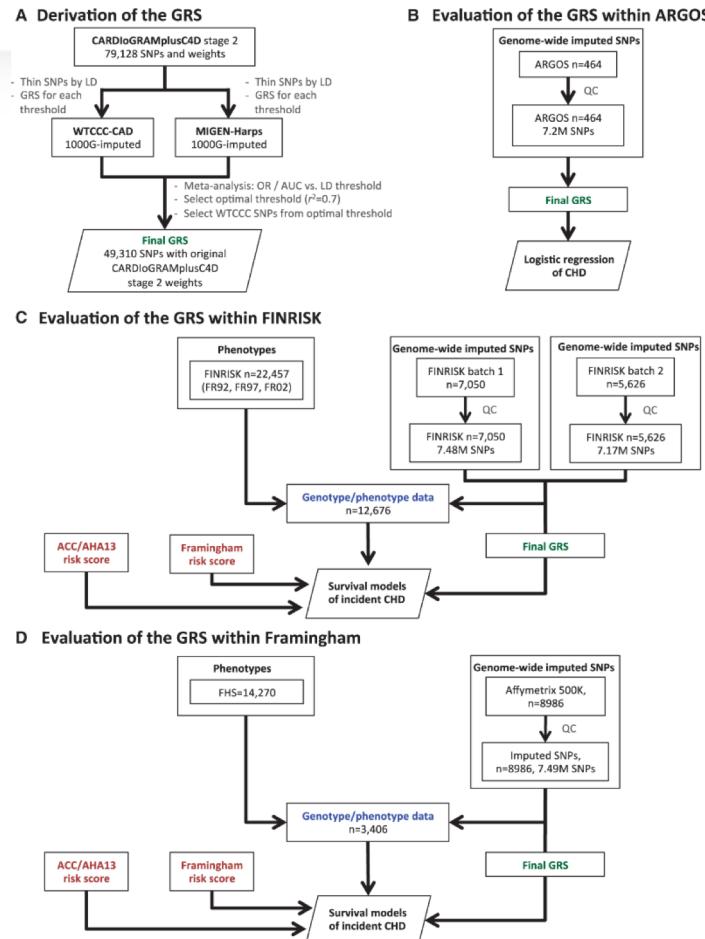
¹Centre for Systems Genomics, School of BioSciences, The University of Melbourne, Parkville, Victoria 3010, Australia; ²Department of Pathology, The University of Melbourne, Parkville, Victoria 3010, Australia; ³National Institute for Health and Welfare, Helsinki FI-00271, Finland; ⁴Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Parkville, Victoria 3010, Australia; ⁵Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki FI-00014, Finland; ⁶Deutsches Herzzentrum München, and Technische Universität München, Munich 80636, Germany; ⁷Deutsches Zentrum für Herz- und Kreislauferkrankungen (DZHK), Partner Site Munich Heart Alliance, Munich 81377, Germany; ⁸Department of Internal Medicine, Erasmus Medical Center, Rotterdam, CA 3000, The Netherlands; ⁹Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, USA; ¹⁰Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA; ¹¹Department of Psychiatry, Psychiatric & Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA; ¹²Department of Cardiovascular Sciences, University of Leicester, BHF Cardiovascular Research Centre, Glenfield Hospital, Groby Rd, Leicester, LE3 9QP, United Kingdom; ¹³National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, United Kingdom; ¹⁴Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom; and ¹⁵Department of Public Health, University of Helsinki, Helsinki FI-00014, Finland

Received 2 May 2016; revised 6 July 2016; accepted 26 August 2016; online publish-ahead-of-print 21 September 2016

See page 3279 for the editorial comment on this article (doi:10.1093/eurheartj/ehw498)



PRS Construction



From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/euroheartj/ehw450

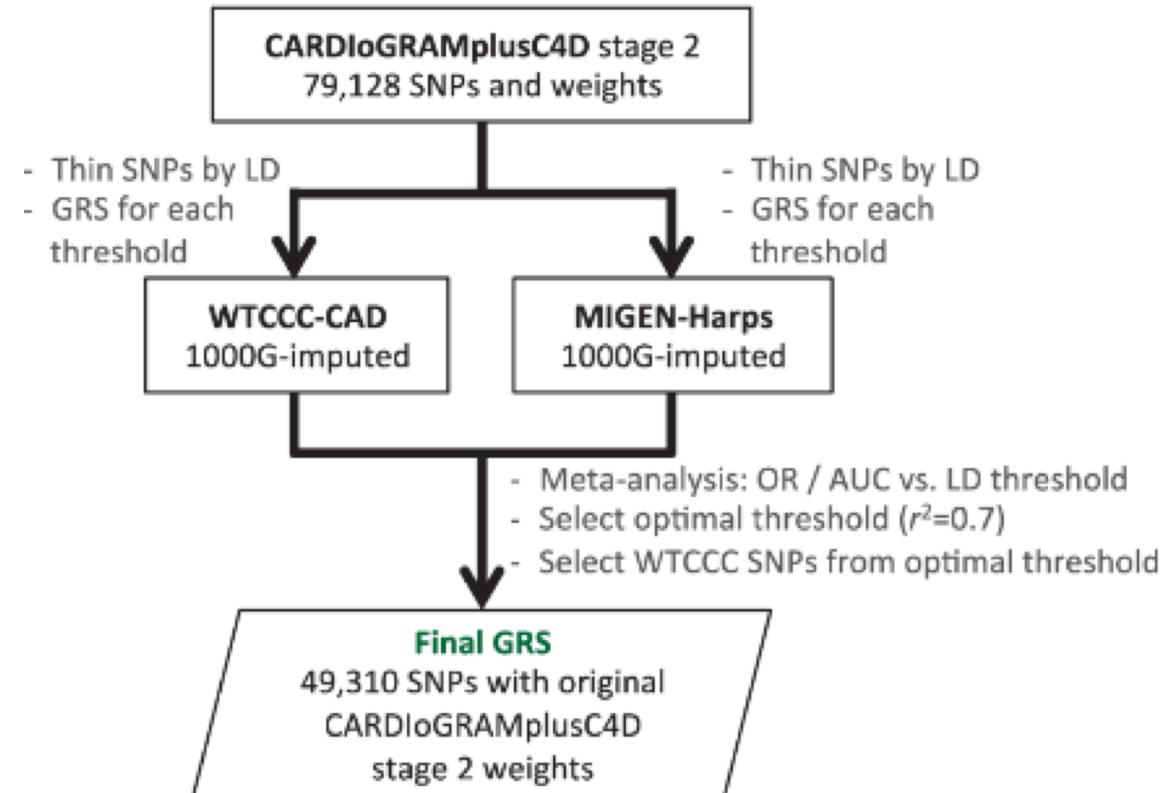
Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



PRS Construction

A Derivation of the GRS



From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/eurheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the

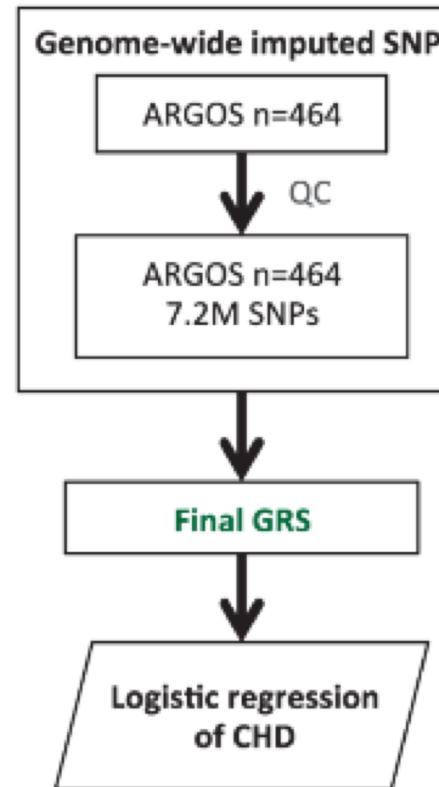
Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



PRS Construction



B Evaluation of the GRS within ARGOS



From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/eurheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

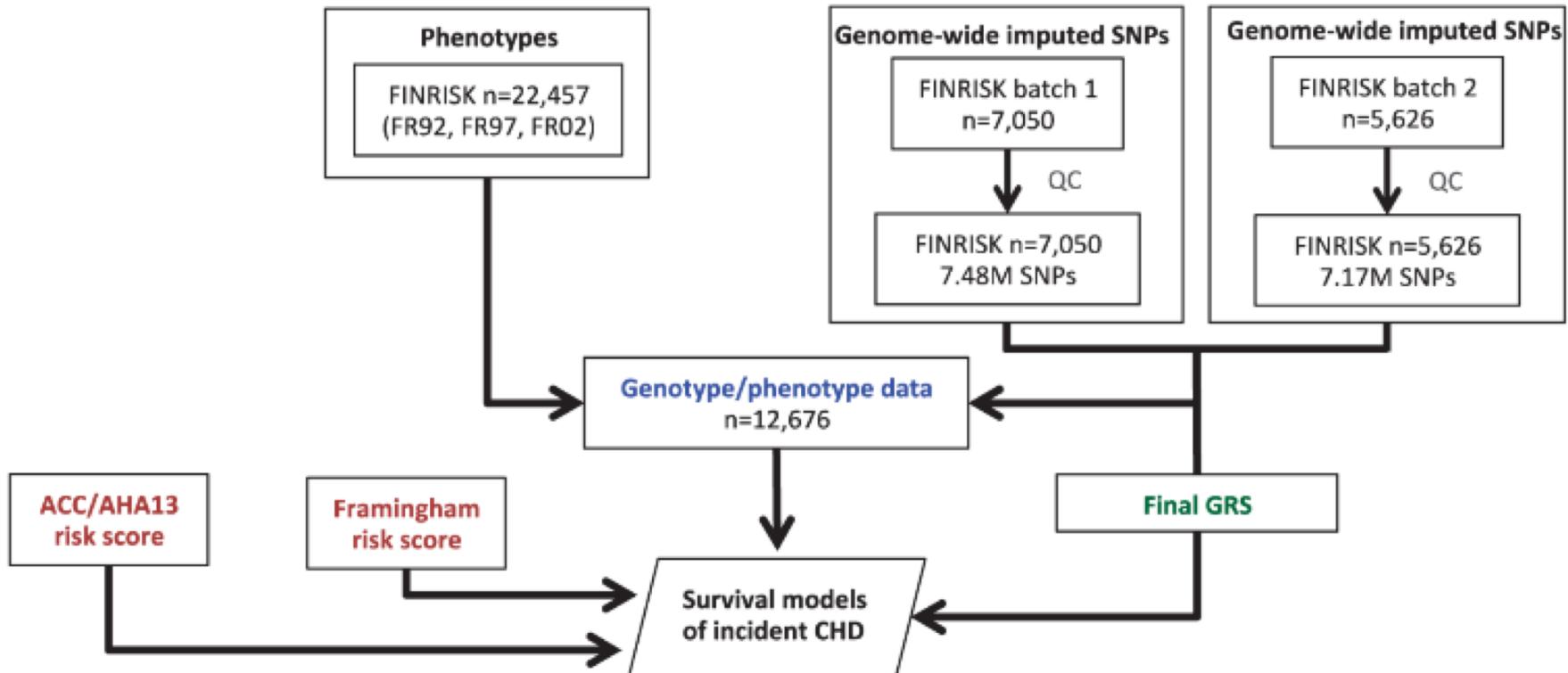
behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



PRS Construction



C Evaluation of the GRS within FINRISK



From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/eurheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

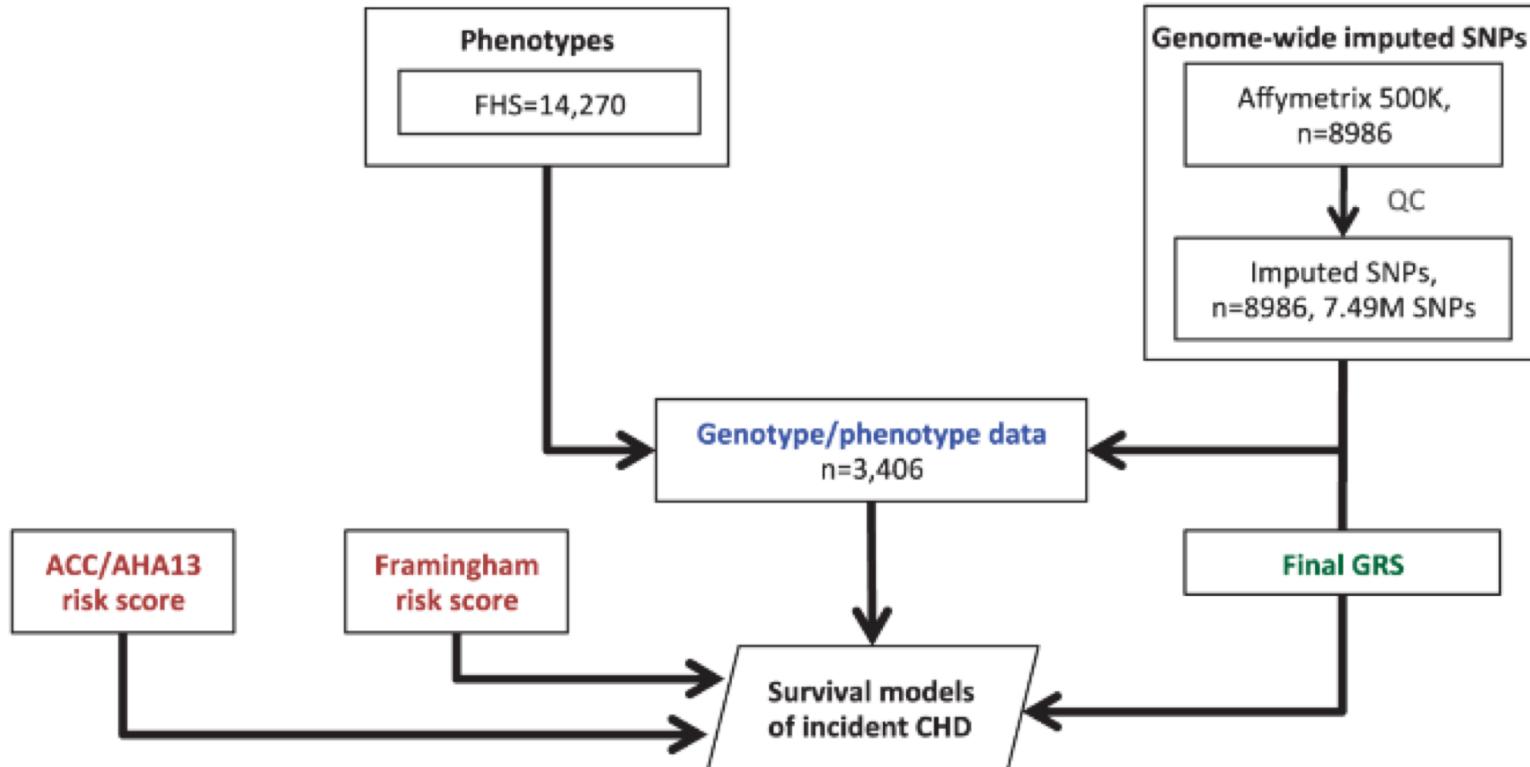
behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



PRS Construction



D Evaluation of the GRS within Framingham



From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/eurheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the

Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



PRS Construction



Table 2

Association of the 49K GRS with incident CHD (binary outcome in logistic regression) in the five studies, per standard deviation of the GRS

Dataset	# Incident CHD/Non-CHD	Odds Ratio (95% CI)
WTCCC-CAD1	1926/2938	1.74 (1.63–1.86)
MIGen-Harps	488/531	1.57 (1.37–1.81)
ARGOS FH	248/216	1.49 (1.21–1.84)
FINRISK	757/11919	1.74 (1.61–1.89)
FHS	587/2819	1.28 (1.17–1.41)

WTCCC-CAD1: adjusted for sex and 5 PCs of the genotypes; MIGen-Harps: adjusted for sex and 5 PCs; ARGOS: adjusted for sex and 5 PCs; FINRISK: adjusted for sex, cohort, east/west, and 5 PCs; FHS: adjusted for sex, cohort, and 5 PCs.

From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/euroheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



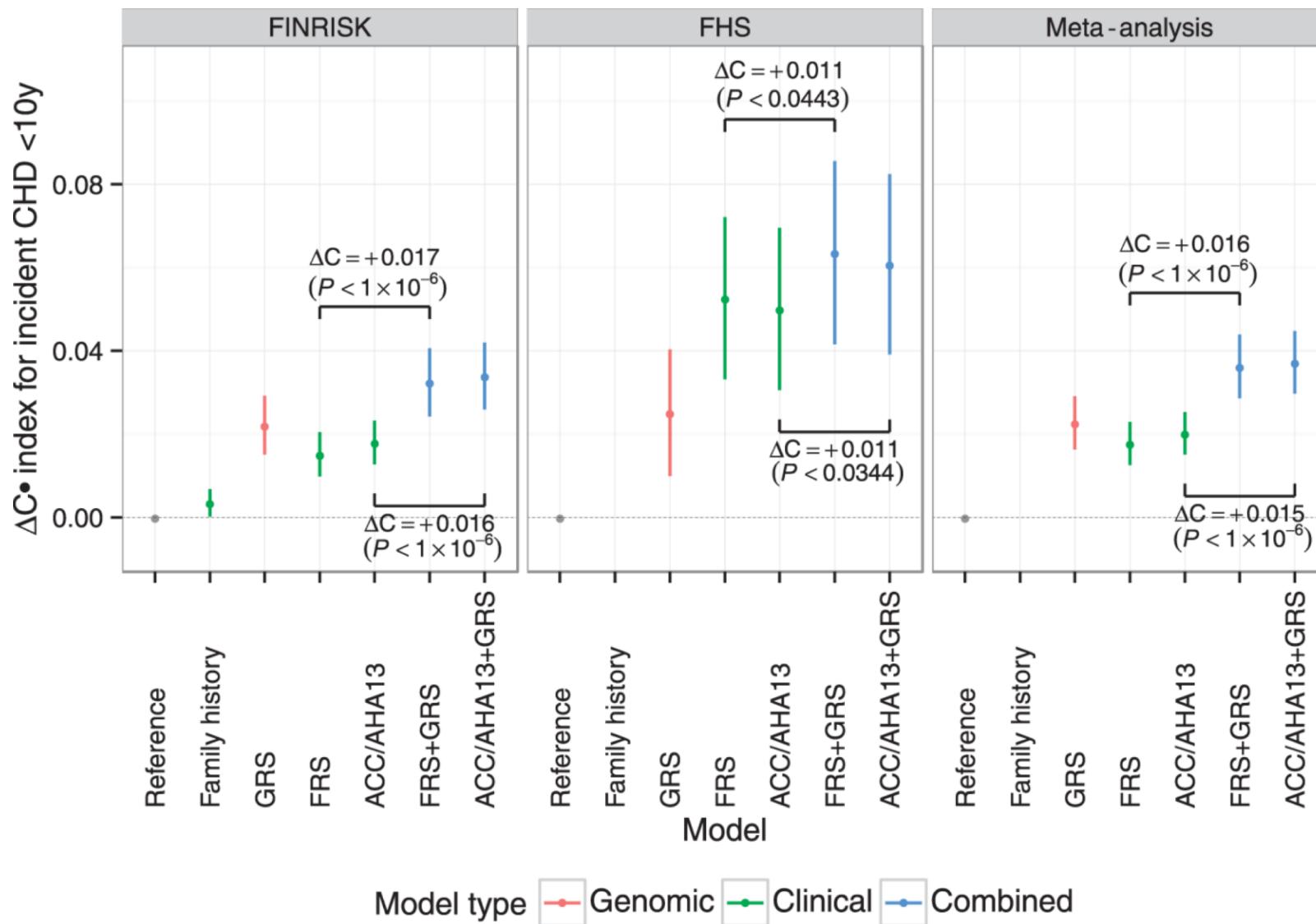
Predictive Power of PRS



- To assess the predictive power of the GRS, we compared its performance in discrimination of time to CHD event (C-index) with that of family history and the FRS and ACC/AHA13 clinical risk scores.

C-statistics

- “concordance” statistic or C-index) is a measure of goodness of fit for binary outcomes in a logistic regression model.
- C-statistic gives the probability a randomly selected patient who experienced an event (e.g. a disease or condition) had a higher risk score than a patient who had not experienced the event.
- It is equal to the area under the Receiver Operating Characteristic (ROC) curve and ranges from 0.5 to 1.



From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/eurheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the

Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 4 Reclassification of incident CHD event risk within 10 years for combined ACC/AHA13 + GRS compared with ACC/AHA13 only, in the FINRISK and FHS cohorts

		FINRISK						FHS					
		ACC/AHA13+GRS						ACC/AHA13+GRS					
		0–7.5%	7.5–10%	10–20%	20–100%	Total	Reclass %	0–7.5%	7.5–10%	10–20%	20–100%	Total	Reclass %
All individuals													
ACC/AHA13	0–7.5%	9,588	211	144	7	9,950	3.6	2,513	78	7	0	2,598	3.3
	7.5–10%	381	176	199	14	770	77.1	112	159	66	1	338	53.0
	10–20%	279	275	755	271	1,580	52.2	7	67	308	32	414	25.6
	20–100%	2	10	127	230	369	37.7	0	0	16	40	56	28.6
	Total	10,250	672	1,225	522	12,699	15.2	2,632	304	397	73	3,406	11.3
All individuals													
FINRISK							FHS						
ACC/AHA13+GRS							ACC/AHA13+GRS						
NRI (categorical)	Total:	0.120	[0.065–0.174];	$P = 1.7 \times 10^{-5}$			Total:	0.068	[−0.014–0.150];	$P = 0.1$			
[95% CI]	NRI for events:	0.097	[0.043–0.151];	$P = 4.52 \times 10^{-4}$			NRI for events:	0.060	[−0.021–0.141];	$P = 0.147$			
	NRI for non-events:	0.023	[0.016–0.030];	$P < 1 \times 10^{-6}$			NRI for non-events:	0.008	[−0.003–0.020];	$P = 0.147$			
NRI (continuous)	Total:	0.356	[0.270–0.442];	$P < 1 \times 10^{-6}$			Total:	0.255	[0.093–0.416];	$P = 0.00197$			
[95% CI]	NRI for events:	0.176	[0.091–0.261];	$P = 4.79 \times 10^{-5}$			NRI for events:	0.160	[0.002–0.318];	$P = 0.047$			
	NRI for non-events:	0.180	[0.164–0.196];	$P < 1 \times 10^{-6}$			NRI for non-events:	0.095	[0.061–0.128];	$P < 1 \times 10^{-6}$			
IDI (continuous)	0.028	[0.021–0.034];	$P < 1 \times 10^{-6}$				0.005	[0.002–0.008];	$P = 0.00184$				
[95% CI]													

In FINRISK, 7 individuals of the 12,676 were excluded in this analysis due to missing clinical measurements.

From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/euroheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the

Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 4 Reclassification of incident CHD event risk within 10 years for cohort ACC/AHA13 only, in the FINRISK and FHS cohorts

	FINRISK					Total	Reclass %	FHS
	ACC/AHA13+GRS				ACC			
	0–7.5%	7.5–10%	10–20%	20–100%				0–7.5%
All individuals								
ACC/AHA13	0–7.5%	9,588	211	144	7	9,950	3.6	2,513
	7.5–10%	381	176	199	14	770	77.1	112
	10–20%	279	275	755	271	1,580	52.2	7
	20–100%	2	10	127	230	369	37.7	0
	Total	10,250	672	1,225	522	12,699	15.2	2,632

From: Genomic prediction of coronary heart disease

Eur Heart J. 2016;37(43):3267-3278. doi:10.1093/eurheartj/ehw450

Eur Heart J | © The Author 2016. Published by Oxford University Press on

behalf of the European Society of Cardiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.



Evaluating the Benefits of PRS



- GeneRisk Study
 - a prospective study that includes about 7,350 randomly selected middle-aged patients from Southern Finland.
- KardioKompassi®
 - Shows 10-year risk for ischemic heart disease.
 - Combines traditional risk factors such as age, sex, cholesterol levels and blood pressure *with a polygenic risk score.*

Abstract no: C01.2 Returning cardiovascular disease risk prediction back to individuals motivates beneficial lifestyle changes : Preliminary results from the GeneRISK study



GeneRisk Study



“Where a patient's overall disease risk was elevated, KardioKompassi advised the participant to contact their doctor in order to discuss how best to reduce it”

-- Elisabeth Widen, MD

GeneRisk Principal Investigator

- Web-based interactive tool
- Motivation for lifestyle change
- *High Polygenic Risk Score for Heart Disease Motivates Patients to Make Lifestyle Changes*

Abstract no: C01.2 Returning cardiovascular disease risk prediction back to individuals motivates beneficial lifestyle changes : Preliminary results from the GeneRISK study



High Risk Group

- Anyone with a combined 10-year CVD risk of more than 10 percent was advised to see a physician,
- ~ 25% of the study cohort
 - 40 percent were smokers
 - 17 percent received statin therapy
 - 12 percent fell into this category because their polygenic risk score had upgraded their clinical risk from a lower-risk category.



Follow-up



- Questionnaire response ~ 5000 participants
 - 90 percent said they had received information that was useful and easy to understand
 - 22 percent found the results to be unexpected
 - 29 percent said the results were of concern to them
- *89 percent said their personal risk information inspired them to make changes to their lifestyle.*



Behavioral changes

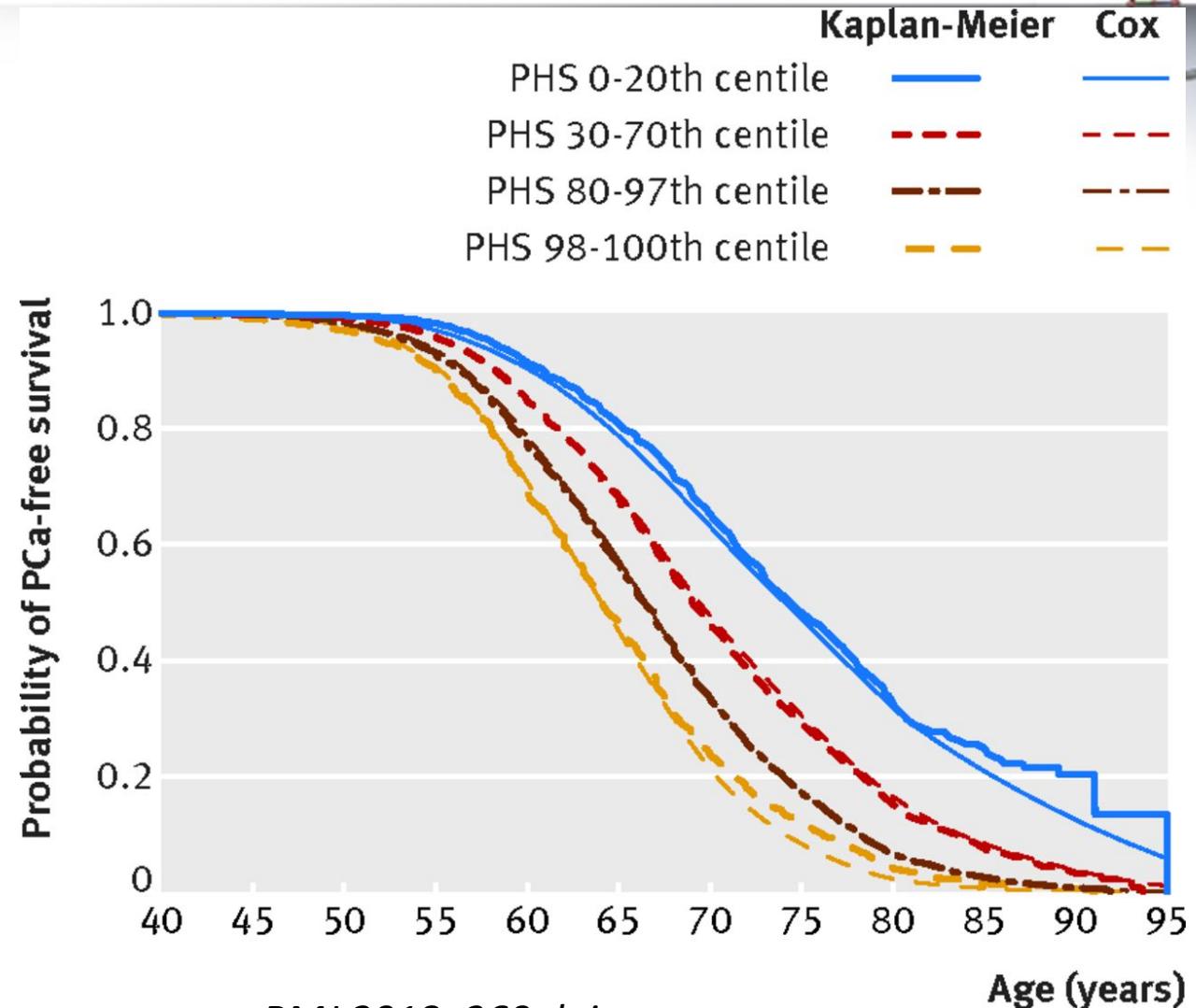
- high-risk and low-risk groups had taken action, but at different rates.
- **36 percent** of those with a CVD risk greater than 10 percent had lost weight, stopped smoking, or seen a physician
- **only 21 percent** of those with a lower CVD risk had done so.



Polygenic Hazard Score & Prostate Cancer



Kaplan-Meier and Cox estimates of prostate cancer-free survival for patients in development set by centile ranges of polygenic hazard score. Centiles are in reference to distribution of score within 11 190 controls aged under 70 in development set. Time of “failure” is age at any diagnosis of prostate cancer. Controls were censored at age of observation. Formal testing of proportionality is described in appendix 1





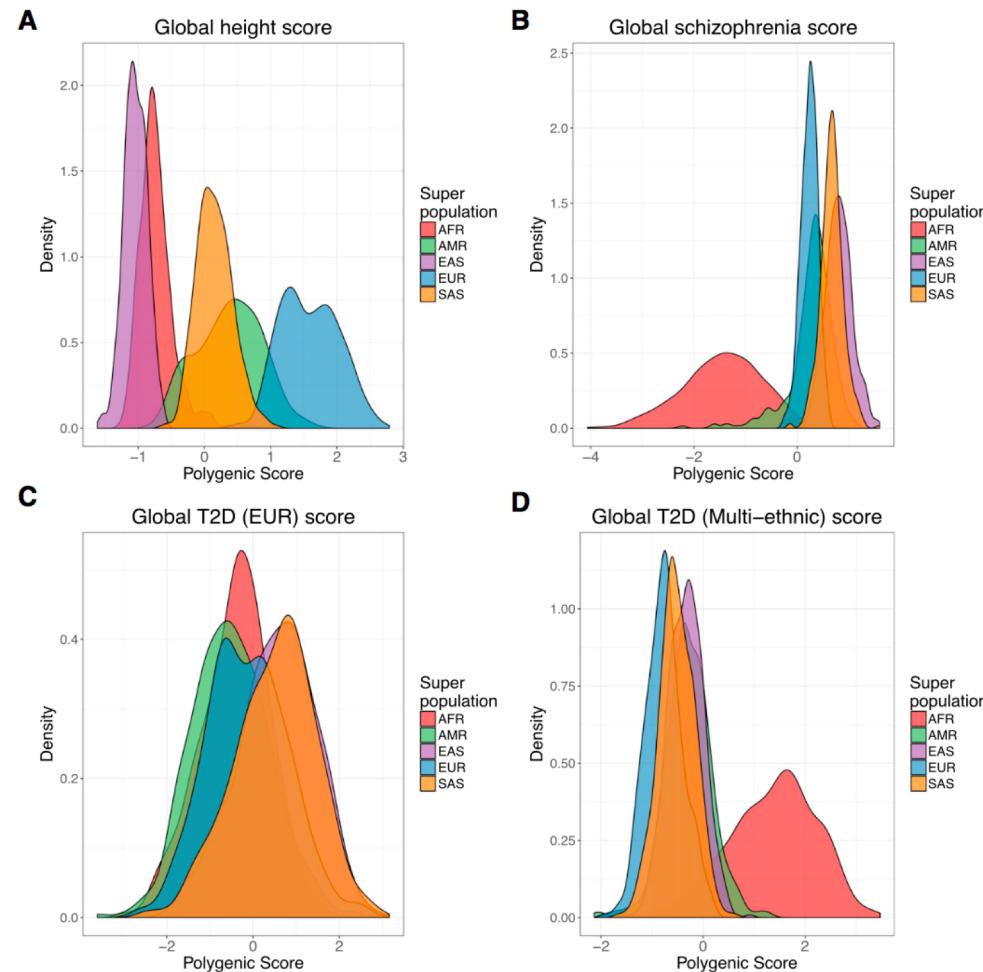
Transferability

- Majority of GWAS were performed in Europeans
- Points to consider:
 - linkage disequilibrium
 - allele frequency
 - genetic architecture





Transferability of PRS

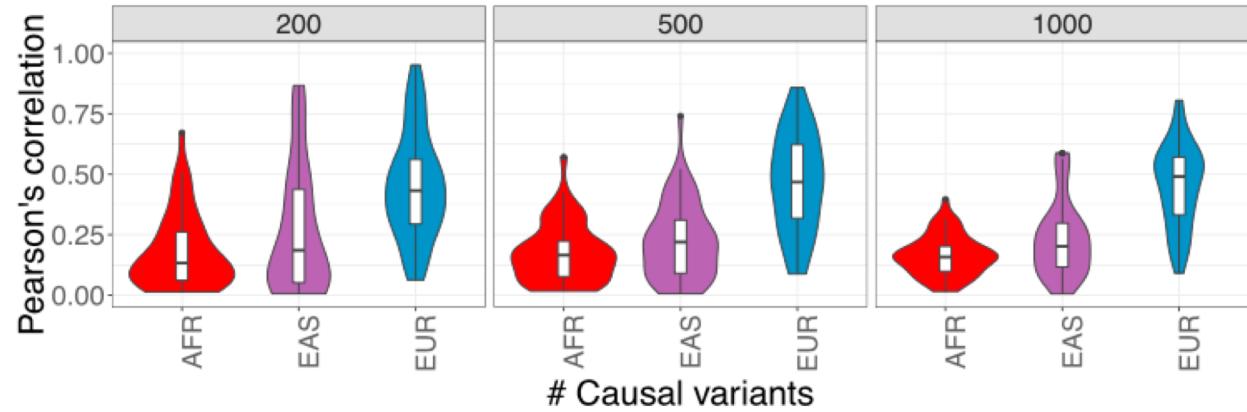




Transferability of PRS

C

Assess risk prediction accuracy.



Data from simulation show that the highest correlation between the prediction accuracy and the reference risk score depends on the source of GWAS



Exercise



- Genome-wide meta-analysis studies identified genetic markers and their effects on migraine as follow (top 3 of 38)

Gene	Marker	OR [95%CI]
FHL5/UFL1	rs67338227	1.09 [1.08–1.11]
near TSPAN2/		
NGF	rs2078371	1.11 [1.09–1.13]
PRDM16	rs10218452	1.11 [1.10–1.13]



Suggested Reading



- 1. Chatterjee N, Shi J, García-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet. 2016 Jul;17(7):392-406. doi: 10.1038/nrg.2016.27. Epub 2016 May 3. Review. PubMed PMID: 27140283; PubMed Central PMCID: PMC6021129.**
- 2.Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet. 2018 Sep;19(9):581-590. doi: 10.1038/s41576-018-0018-x. Review. PubMed PMID: 29789686.**



Solution

- Using weighted sum method to combine the odds ratio
- $OR = \exp(\beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3)$
- Since we were given that $\exp(\beta_1) = 1.09$; $\exp(\beta_2) = 1.11$; $\exp(\beta_3) = 1.11$
- $OR =$
 $= \exp[\log(1.09)*X_1 + \log(1.11)*X_2 + \log(1.11)*X_3]$
 $= \exp[\log(1.09)*0 + \log(1.11)*1 + \log(1.11)*1]$
 $= 1.2321$